AMERICAN PSYCHOLOGICAL ASSOCIATION

Emotion

ISSN: 1528-3542

2024, Vol. 24, No. 7, 1663–1675
https://doi.org/10.1037/emo0001359

Emotional Scenes as Context in Emotional Expression Recognition: The Role of Emotion or Valence Match

Lachlan Bryce[1], Georgia Mika[2], Belinda M. Craig[3], Ursula Hess[4], and Ottmar V. Lipp[1, 2]
[1] School of Population Health, Curtin University
[2] School of Psychology and Counselling, Queensland University of Technology
[3] Faculty of Heath Sciences and Medicine, Bond University
[4] Department of Psychology, Humboldt University

Emotion recognition is influenced by contextual information such as social category cues or background scenes. However, past studies yielded mixed findings regarding whether broad valence or specific emotion matches drive context effects and how multiple sources of contextual information may influence emotion recognition. To address these questions, participants were asked to categorize expressions on male and female faces posing happiness and anger and happiness and fear on pleasant and fearful backgrounds (Experiment 1, conducted in 2019), fearful and disgusted expressions on fear and disgust eliciting backgrounds (Experiment 2, conducted in 2022), and fearful and sad expressions on fear and sadness eliciting backgrounds (Experiment 3, conducted in 2022). In Experiment 1 (where stimuli varied in valence), a broad valence match effect was observed. Faster recognition of happiness than fear *and* anger was more pronounced in pleasant compared to fearful scenes. In Experiments 2 and 3 (where stimuli were negative in valence), specific emotion match effects were observed. Faster recognition occurred when expression and background were emotionally congruent. In Experiments 1 and 3, poser sex independently moderated emotional expression recognition speed. These results suggest that the effect of emotional scenes on facial emotion recognition is mediated by a match in valence when broad valence is task-relevant. Specific emotion matches drive context effects when participants categorize expressions of a single valence. Looking at the influence of background contexts and poser sex together suggests that these two sources of contextual information have an independent rather than an interactive influence on emotional expression recognition speed.

Keywords: context, emotion perception, facial expression recognition, happy face advantage, social category cues

Supplemental materials: https://doi.org/10.1037/emo0001359.supp

Facial expressions of emotion are important in regulating social interactions in everyday life (Niedenthal & Brauer, 2012). Happy smiles or sad frowns signal that the expresser can be safely approached, whereas an expression of anger suggests someone to be avoided. However, like any social signal, facial expressions do not occur in a vacuum. Their meaning and the manner in which they are perceived are influenced by contextual information. Information that contextualizes emotional expressions can be derived from a poser's face (Hess et al., 2009; Hugenberg & Sczesny, 2006; Lipp et al., 2015) or body posture (Meeren et al., 2005), from the physical environment in which the posers are encountered (Kastendieck et al., 2021; Leppänen & Hietanen, 2003; Righart & de Gelder, 2008a, 2008b), from a poser's voice (de Gelder & Vroomen, 2000), or from knowledge we have about the poser (Diéguez-Risco et al., 2013;

This article was published Online First June 13, 2024.
Philippe G. Schyns served as action editor.
Ursula Hess https://orcid.org/0000-0002-6793-7282
Ottmar V. Lipp https://orcid.org/0000-0001-6734-8608
This research was supported under the Australian Research Council's Discovery Projects funding scheme (Grants DP150101540 and DP220100878 awarded to Ottmar V. Lipp).
The code used to run the experiments and the data reported in this article are available at https://osf.io/4h6v7/ (Bryce et al., 2023). The experiments were not preregistered.

Lachlan Bryce played a lead role in investigation and a supporting role in conceptualization, data curation, and writing–review and editing. Georgia Mika played a lead role in investigation and a supporting role in conceptualization, data curation, and writing–review and editing. Belinda M. Craig played a lead role in formal analysis, a supporting role in writing–review and editing, and an equal role in conceptualization, data curation, and funding acquisition. Ursula Hess played a lead role in writing–review and editing and an equal role in conceptualization, data curation, formal analysis, and funding acquisition. Ottmar V. Lipp played a lead role in conceptualization, data curation, formal analysis, funding acquisition, supervision, and writing–original draft and an equal role in methodology.
Correspondence concerning this article should be addressed to Ottmar V. Lipp, School of Psychology and Counselling, Queensland University of Technology, Kelvin Grove, QLD 4059, Australia. Email: Ottmar.Lipp@qut.edu.au

Lindeberg et al., 2019; for a review, see Wieser & Brosch, 2012). Given the wealth of information about the factors that can affect expression perception, it is important to determine the mechanisms that mediate the influence of contextual cues on facial emotion recognition and the manner in which different sources of contextual information interact. The current research was designed to address some of these issues.

Contexts can affect the perception of and the reaction to emotional expressions in different ways. On one hand, situational context or knowledge about an expresser's likely emotional reactions can allow observers to disambiguate ambiguous expressions using perspective-taking (Hess & Hareli, 2016). Conversely, on a perceptual level, facial morphology can facilitate the perception of some emotions and hinder the perception of others (Hess et al., 2009). Also on the perceptual level, emotion recognition speed may be facilitated when context and emotional expression match by making the relevant emotion categories more easily accessible (Zhang et al., 2013). To assess the effects of different sources of contextual information, we investigated the effects of both social category cues and context cues on emotion recognition speed.

Past research from our laboratory and others has investigated the effects of social category cues that are easily derived from a face, such as sex, age, or ethnicity, on emotion recognition speed (Bijlstra et al., 2010; Craig & Lipp, 2018; Hugenberg, 2005; Hugenberg & Sczesny, 2006; Lipp et al., 2015). This research confirmed that expressions of happiness are recognized faster than expressions of anger or other negative emotions (Craig, Koch, & Lipp, 2017) on faces that are female, young, or from the ethnic ingroup, but not on faces that are male, old, or from the ethnic outgroup (Craig & Lipp, 2018; Craig, Zhang, & Lipp, 2017; Lipp et al., 2015). Research has proposed a number of mechanisms for the influence of social category information on emotion recognition, including the role of facial structure (e.g., Becker et al., 2007; Hess et al., 2009; see Craig & Lee, 2020 for a review), stereotypes about social groups (Bijlstra et al., 2010, 2019), and broad social evaluations (Hugenberg, 2005; Hugenberg & Sczesny, 2006).

The effect of facial social category cues on this happy advantage is most comprehensively explained within Hugenberg's (2005) evaluative congruence account, which stipulates that the perception of a relatively positively evaluated social category cue will facilitate the recognition of a positively valenced expression. This has been concluded as the same categorization patterns have been observed in tasks where participants categorize expressions (such as anger) that are both evaluatively and stereotypically congruent with a particular social group; for example, men are evaluated as relatively negative compared to women (Eagly et al., 1991) and are stereotyped as more aggressive (Plant et al., 2000), as well as evaluatively congruent but stereotypically incongruent emotions like sadness; for example, women are relatively positively evaluated (Eagly et al., 1991) but are more likely to be associated with sadness than men (Plant et al., 2000). Finding faster categorization of happiness than anger as well as happiness than sadness on female versus male faces aligns with the evaluative congruence account. Further evidence for the role of evaluations over facial structure comes from studies showing that emotion categorization of the same faces depends on the social context in which they are encountered. Happiness is recognized faster on Caucasian male faces when presented among African American male faces (Lipp et al., 2015) or Chinese male faces (Craig, Zhang, & Lipp, 2017) but not when the same Caucasian male faces are presented among female faces (where the male category becomes relatively negatively evaluated).

Studies have found that sex or racial stereotypes (Bijlstra et al., 2010) can also influence emotion recognition, but only in contexts where broad positive–negative valence is not task-relevant. For example, the effect of stereotypes on emotion recognition has been observed in tasks where participants categorize sadness versus anger (two negatively valenced expressions) on male and female faces. Participants are faster to categorize sadness than anger on female faces and anger than sadness on male faces (Bijlstra et al., 2010). These studies demonstrate that, where broad positive–negative valence is task-relevant, social category valence takes precedence, but stereotypic associations can also influence the speed of emotion recognition when valence is not primed by the task.

Beyond the face, the influence of environmental cues on emotion recognition speed has been investigated by immersing participants in environments that differ in valence (Leppänen & Hietanen, 2003) or by presenting the target faces embedded in emotion-provoking scenes (Hietanen & Astikainen, 2013; Milanak & Berenbaum, 2014; Righart & de Gelder, 2008a, 2008b; Xu et al., 2015, 2017; Zhang et al., 2013). These studies have not only provided evidence that environmental cues influence emotion perception but also have generated a number of questions as to the mechanism/s that may underlie the observed effects. As outlined in more detail in the following section, most previous studies seem designed from a discrete emotion perspective, utilizing emotional scenes and facial expressions that represent the same discrete emotions, whereas others are situated in a dimensional framework, selecting background stimuli and facial expressions that match in valence but not necessarily in discrete emotion. Our work on the effects of facial social category cues and the evaluative congruence account would support the latter approach, but a direct empirical test of which account best encompasses context effects is currently lacking.

Emotion Versus Valence Match

A common approach to assessing context effects on emotion recognition is to embed facial expressions in scenes with matching or mismatching emotional content (i.e., a trash heap together with a disgust face vs. a happy face). Using such a design, Righart and de Gelder (2008a) found that happiness was recognized faster when presented in happy or neutral scenes than in fear scenes, but recognition of fear was not affected by scene emotion. In a subsequent study, Righart and de Gelder (2008b) broadened the range of expressions, presenting happy, fearful, and disgusted faces in scenes aimed at eliciting these emotions. Using a three alternative forced choice task, Righart and de Gelder (2008b) found faster recognition of disgust in disgust and happy scenes than in fear scenes, faster recognition of fear in fear and happy scenes than in disgust scenes, and faster recognition of happiness in happy than in fear scenes. Adding a secondary task to increase task load in a second experiment, Righart and de Gelder (2008b) obtained a stronger effect of specific emotion match in that emotion recognition was faster for all three expressions in scenes that matched the expressed emotion relative to the two scenes that mismatched. Milanak and Berenbaum (2014) reported a similar emotion match effect for the recognition of sadness, fear, and happiness, but not for disgust.

Zhang et al. (2013) varied whether expressions and scenes were presented simultaneously or sequentially and found in the simultaneous condition that recognition of disgust and happiness, but not fear, was faster in scenes that matched the emotion than in scenes that mismatched, yet no background effect emerged for accuracy. In the sequential condition, recognition of fear and disgust, but not happiness, was more accurate after matching scenes, but no background effect emerged for recognition time. Xu et al. (2015) also failed to find effects of emotion-matched scenes on the speed or accuracy of recognizing fear or happiness. By contrast, an effect of scene emotion on facial emotion recognition was observed by Xu et al. (2017), who found that recognition of happy and fearful faces was both more accurate and faster on emotion matching than on mismatching scenes. These findings suggest that a match between the discrete emotion elicited by a background scene and the expressed emotion can facilitate emotional expression recognition; however, it is also evident that the effect is inconsistent with regard to which emotions were affected and whether the effect pertained to accuracy or speed.

The above-cited research matched scenes and expressions by discrete emotion category. That scene valence rather than discreet emotion category can affect emotion recognition was shown by Hietanen and Astikainen (2013). These authors presented pleasant or unpleasant scenes from the International Affective Picture System (IAPS), followed by happy or sad faces. The scenes used in this study were selected based on their pleasantness and not on the discrete emotion they conveyed. Emotion recognition was more accurate and faster in the valence match than in the valence mismatch conditions (marginally for recognition speed of sadness). Thus, the results of Hietanen and Astikainen (2013) suggest that valence match of background and facial expression may be sufficient to support facilitated emotion recognition.

It should be noted that the interpretation of some studies that utilized emotion matching between background scenes and expressions is complicated by the fact that performance in the match condition was compared to a mismatch condition created by collapsing across all possible mismatches (e.g., happy and fear scenes vs. disgust scenes for disgust recognition). This potentially confounds the effects of scene valence with the effects of scene emotion, as, depending on the combination of expressions and background scenes used in the task, the influence of different combinations of positively and negatively valenced backgrounds is combined to create the "incongruent" condition (Milanak & Berenbaum, 2014; Zhang et al., 2013).

Current Project

The above review of past work on emotional scene context effects highlights a number of inconsistencies that require clarification. First, the nature of the relationship between context emotion and emotional expression that will lead to facilitated emotion recognition is unclear. Most past studies seem to be designed assuming that the scene depicted in the context has to match the discrete emotion expressed on the face; however, strong evidence for emotion specificity seems limited; that is, it was evident in Xu et al. (2017) and the second experiment by Righart and de Gelder (2008b), but not in their first, and not in Milanak & Berenbaum (2014) or Righart and de Gelder (2008a). Hietanen and Astikainen (2013), on the other hand, found that a valence match was sufficient to enhance recognition of happiness and (marginally) sadness. It should also be noted that the majority of the past studies employed small samples of 21 participants or fewer (see Milanak & Berenbaum, 2014, for an exception), which may have limited their power.

Finally, it remains unclear whether and how multiple sources of contextual information presented simultaneously can affect emotion recognition speed. The majority of the past studies using emotional scene backgrounds have employed male and female faces but did not examine emotion recognition speed as a function of poser sex. This is surprising as the influence of face sex on emotion recognition speed is well documented (Becker et al., 2007; Craig & Lipp, 2017; Hugenberg & Sczesny, 2006). For instance, happiness is very consistently recognized more quickly and accurately on the faces of women than men. The reverse has been shown for anger (Becker et al., 2007; Hess et al., 2009), although anger superiority for men's faces is not always observed in speeded emotion recognition tasks (Hugenberg & Sczesny, 2006; Lipp et al., 2015).

More recently, we (and others) have investigated how multiple sources of contextual (social) information (e.g., poser race and sex, poser sex and age) combine to influence emotion perception (e.g., Craig & Lipp, 2018; Lindeberg et al., 2019; Smith et al., 2017). In the majority of previous studies, the multiple category cues have an additive rather than an interactive influence on emotion recognition speed and accuracy. For example, when the sex and the race of the face are both varied, a Poser Sex × Facial Expression interaction and a Poser Race × Facial Expression interaction were both observed in the absence of a Poser Sex × Poser Race × Facial Expression interaction. Investigating how two different sources of contextual information on and outside the face may interact or combine (in this case, focusing on poser sex and background scenes) is an extension of this broader avenue of research. It remains unknown whether and how the effects of emotional scene contexts and facial social category cues combine.

The present research was designed to extend our understanding of the manner in which emotional scenes and facial sex cues affect the recognition of emotional expressions. Experiment 1 assessed whether the influence of background scenes on emotion recognition speed is driven by a match of discrete emotion elicited by a scene or broad positive–negative valence elicited by the scene. To achieve this, participants were asked to recognize either expressions of happiness and fear or of happiness and anger presented in scenes associated with happiness or fear across two tasks. If the facilitation of emotion recognition is limited to discrete emotion matches, then faster responding when face and background match should be evident in the fear task but not in the anger task. If it is scene valence match that is sufficient to facilitate emotion recognition, then facilitated recognition when expression and scene are evaluatively congruent should be evident in both fear and anger tasks.

To date, studies investigating scene context effects have included at least one positive and one (sometimes multiple) negative expression, possibly increasing the salience/relevance of broad positive–negative evaluations (Bijlstra et al., 2010). As such, Experiments 2 and 3 aimed to further examine the role of specific emotion match in single-valence tasks (i.e., where only negatively valenced expressions and scenes are encountered) to determine whether specific emotion match patterns are observed once the salience of broad valence is removed. Participants were asked to categorize disgusted and fearful expressions on disgust and fear eliciting backgrounds (Experiment 2) or fearful and sad expressions on fear and sadness eliciting backgrounds (Experiment 3). If specific emotion match effects are present when valence and specific emotions are unconfounded in match and mismatch trials in the task design, participants should be faster to categorize expressions when face and background are congruent with regards to a specific emotion than when face and background are incongruent. In all experiments, we present male and female posers to assess the combined effects of emotional scenes and facial sex cues on emotion recognition.

Experiment 1

Method

Participants

A power analysis (G-power; Faul et al., 2007) based on the averaged effect sizes reported by Righart and de Gelder (2008b) for Scene × Expression interactions ($\eta_p^2 = .35$) suggested that to achieve a power of .8 at α .05 to find a significant difference among 16 measurements (2 × 2 × 2 × 2 design; option "as in Cohen") would require 36 participants. However, as published effect size estimates are often inflated, we decided to oversample by approximately 40%. Fifty-two undergraduate students provided informed consent and volunteered participation in exchange for course credit. After exclusion of two participants for excessive errors (>20%) on one of the categorization tasks, the final sample comprised 50 participants (33 women and 17 men; age range: 17–28 years; $M_{\text{age}} = 20.28$; $SD_{\text{age}} = 2.12$). All participants had normal or corrected-to-normal vision. Participants self-reported their ethnicity as Caucasian (33), Asian (11), African (2), Indigenous Australian (1), or other (3). The experimental protocol was approved by the Curtin University Institutional Ethics Review Board (HRE2017-0348). Experiment 1 was conducted in 2019.

Apparatus, Materials, and Procedure

Participants completed the experiment in groups of up to four using computers with a screen size of 1,920 × 1,080 pixels and a refresh rate of 120 Hz. The tasks were controlled using Millisecond's Inquisit 4 (Inquisit, 2015). Pictures of 16 Caucasian posers, eight men and eight women, were drawn from the NimStim face database (Tottenham et al., 2009; Poses AN_O and HA_O, FE_O of models 1, 2, 3, 5, 6, 20, 21, 23, 24, 28) and the Montreal Set of Facial Displays of Emotion (Beaupré & Hess, 2005; Poses 1, 2, and 4 of models 20, 22, 23, 25, 27, 28). Face images were edited to show only the face, converted to grayscale, and resized to 187 × 240 pixels. Based on the ratings provided by Mikels et al. (2005), 16 color images were drawn from the IAPS (Lang et al., 2008) to act as emotional background scenes (fearful: 1,019, 1,052, 1,113, 1,200, 1,300, 1,930, 1,931, 9,600; pleasant:[1] 1,440, 1,460, 1,463, 1,710, 5,480, 5,760, 5,814, 7,502). Images were presented in color at a size of 1,024 × 768 pixels. During the recognition tasks, faces were presented centered on the image backgrounds.

The experimental procedure comprised two expression categorization tasks, one with happy and fearful and one with happy and angry expressions, each completed twice, yielding two blocks per task. Participants were instructed to look at the screen and to press the "S" or "L" keys on a QWERTY keyboard as quickly as possible to indicate whether the expression on the displayed face was happy or fearful/angry. Response mapping was held constant for a participant but counterbalanced across participants. During each task, the question "Is the face happy or fearful/angry?" was displayed above the location of the image/face compound, and the reminders of the response buttons (e.g., "Press S for Angry" and "Press L for Happy") were presented in the left or right corners of the screen below the image/face compound. Each block comprised 12 practice trials with response feedback and 128 test trials without feedback, yielding 256 trials per task. Trials were separated by a 1,000-ms intertrial interval during which the text remained on screen and consisted of the presentation of a centered fixation cross for 500 ms, followed by an image/face compound with simultaneous onset presented for 3,000 ms, or until a response was made. Each task employed eight different background images (four fearful and four pleasant) and eight posers (four females and four males). Task sequence was alternated and counterbalanced across participants, as was the allocation of background images and posers to the tasks and the match of response button to expression.

The emotion recognition tasks were followed by a rating task in which participants rated the pleasantness of the background images on a 7-point Likert scale and selected a label that best described the emotion evoked by the scene. Labels provided were sadness, contentment, happiness, fear, surprise, anger, and disgust.

Design and Analyses

Prior to analysis, data from the two blocks of each task were combined. Trials with incorrect responses, recognition times faster than 100 ms, or recognition times that differed from a participant's mean by more than 3 *SD*s were classed as errors and excluded from the analysis of recognition times. Data from two participants with more than 20% errors were excluded from analyses. Separate 2 (Task: fear vs. anger) × 2 (Background Emotion: pleasant vs. fear) × 2 (Poser Sex: male vs. female) × 2 (Facial expression: happy vs. angry/fearful) repeated measures analyses of variance (ANOVAs) were conducted on response times and error data, respectively (across experiments, the analysis of error data did not provide additional information. In the interest of brevity, error rate analyses are reported in Supplemental Material). Pairwise comparisons were conducted for all significant interactions. Preliminary analyses revealed no effect of task sequence or participant sex on task performance; hence, these factors are not considered further.

Pleasantness ratings for the fearful and pleasant scenes were averaged within category and subjected to a pairwise *t* test. Emotion label assignments were summed across participants and averaged across images such that a score of 50 would indicate that all participants chose a particular label for a particular set of images (see Supplemental Material for labeling of individual images).

In addition, Jeffreys-Zellner-Siow Bayes factor ANOVAs with default priors were performed using the software package JASP (JASP Team, 2024; Morey et al., 2016; Rouder et al., 2012). These analyses provide information on how well different statistical models, including each main effect and interaction, as well as all combinations of these effects, fit the observed data by providing a

[1] Mikels et al. (2005) did not include the label "happiness" in their study on emotional categories for positive IAPS pictures but used the labels "awe," "excitement," "contentment," and "amusement." To reflect this and to enhance the distinction between emotional scenes and expression, we decided to refer to these pictures as "pleasant."

Bayes factor (BF) for each model. The largest BF indicates the model that best fits the data. BFs can also be interpreted using conventions to indicate whether there is evidence for a given model (Kass & Raftery, 1995). A BF_{10} between 1 and 3 suggests anecdotal evidence, between 3 and 20 suggests positive evidence, between 20 and 150 provides strong evidence, and >150 suggests very strong evidence for a model. BFs are also transitive (one BF can be divided by the other to create a BF comparing the two models). This means that two models can be compared to determine whether there is greater evidence for one model over another given the observed data.

Background Scene Ratings and Labeling

Participants rated the fear-eliciting background images as less pleasant than the pleasant backgrounds, $M = 1.99$, $SD = 0.62$ versus $M = 5.97$, $SD = 0.44$, $t(49) = 32.34$; $p < .001$; $d_z = 4.59$. The eight fear-related images were predominantly labeled as evoking fear ($M = 33.13$, $SD = 12.53$), with two of the images also labeled as evoking disgust (IAPS 1,019: 34 ratings) or sadness (IAPS 9,600: 31 ratings). Overall, fearful images were not consistently labeled as evoking anger ($M = 1.13$, $SD = 1.73$), sadness ($M = 4.00$, $SD = 10.92$), disgust ($M = 8.25$, $SD = 11.42$), surprise ($M = 1.38$, $SD = 1.06$), happiness ($M = 0.75$, $SD = 0.71$), or contentment ($M = 1.13$, $SD = 1.13$). Pleasant images were predominately labeled as evoking happiness ($M = 27.38$, $SD = 13.57$) or contentment ($M = 17.00$, $SD = 10.84$), but not fear ($M = 1.00$, $SD = 1.85$), anger ($M = 1.13$, $SD = 0.35$), sadness ($M = 1.38$, $SD = 1.69$), surprise ($M = 3.00$, $SD = 4.57$), or disgust ($M = 0.13$, $SD = 0.35$). These results are consistent with the data provided by Mikels et al. (2005).

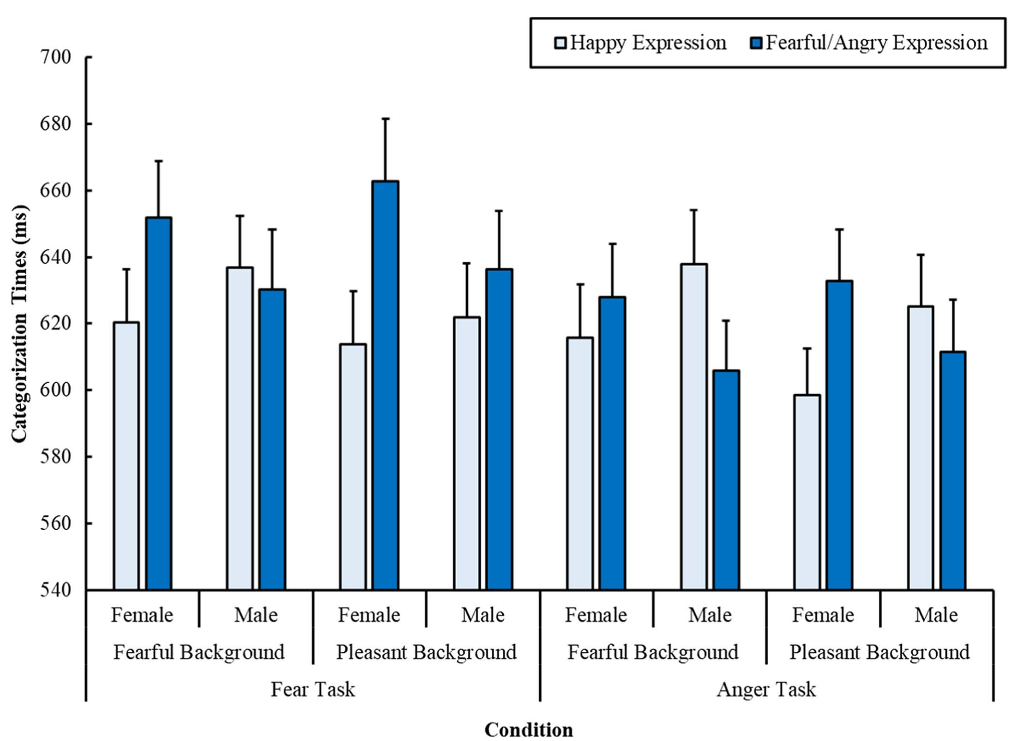
Emotion Recognition Speed

As shown in Figure 1, poser sex and background scene valence affected the recognition of happy and angry/fearful expressions, such that happiness was recognized faster when expressed by women compared to men and anger/fear was recognized faster when shown by men rather than women. Overall, happiness seemed to be recognized faster, and this happy face advantage was more pronounced for pleasant compared to fearful scenes. Most importantly, the overall pattern of results was remarkably consistent across tasks.

The analysis confirmed this interpretation, yielding a main effect for facial expression, $F(1, 49) = 5.39$, $p = .024$, $\eta_p^2 = .10$, and Task × Facial Expression, $F(1, 49) = 11.66$, $p = .001$, $\eta_p^2 = .19$, Background Emotion × Facial Expression, $F(1, 49) = 10.18$, $p = .002$, $\eta_p^2 = .17$, and Poser Sex × Facial Expression interactions, $F(1, 49) = 47.82$, $p < .001$, $\eta_p^2 = .49$. No other interactions involving the task factor were significant, all $F(1, 49) < 0.88$, $p > .353$, $\eta_p^2 < .02$. The Task × Facial Expression interaction reflected that angry expressions were recognized faster than fearful expressions, $F(1, 49) = 6.69$, $p = .013$, $\eta_p^2 = .12$, whereas there was no difference in recognizing happy

Figure 1

Recognition Times for Happy or Fearful/Angry Expressions on Women's and Men's Faces as a Function of Background Scene Emotion and Task in Experiment 1 (Error Bars Represent Standard Errors of the Mean)



Note. See the online article for the color version of this figure.

expressions between tasks, $F(1, 49) = 0.23$, $p = .635$, $\eta_p^2 = .01$. The Background Emotion × Facial Expression interaction was due to faster recognition of happiness in pleasant compared to fear backgrounds, $F(1, 49) = 9.93$, $p = .003$, $\eta_p^2 = .17$, whereas there was only a trend toward faster recognition of anger or fear in fear backgrounds, $F(1, 49) = 3.20$, $p = .080$, $\eta_p^2 = .06$. Finally, the Poser Sex × Facial Expression interaction reflected faster recognition of happiness than anger and fear shown by women, $F(1, 49)=32.67$, $p < .001$, $\eta_p^2 = .40$, but not men, $F(1, 49) = 2.71$, $p = .106$, $\eta_p^2 = .05$.

Although the Task × Background Emotion × Facial Expression interaction was not significant on the preset level, $F(1, 49) = 0.01$, $p = .915$, $\eta_p^2 < .01$, we tested for Background Emotion × Facial Expression interactions in each task, as these analyses are critical to the research question under investigation. Consistent with the role of broad valence, the Background Emotion × Facial Expression interactions were significant for the fear, $F(1, 49) = 5.30$, $p = .026$, $\eta_p^2 = .10$, and for the anger task, $F(1, 49) = 10.72$, $p = .002$, $\eta_p^2 = .18$.

Supplementary Bayesian Analyses

The largest BF emerged for the model, including the main effects of task, poser sex, and facial expression, as well as Task × Facial Expression and Poser Sex × Facial Expression interactions $BF_{10} = 2.08 \times 10^9$. More consistent with the frequentist analyses, the model with the next largest effect size also included the Background Emotion × Facial Expression interaction (model: main effects of task, background emotion, sex, facial expression, and interactions of Task × Facial Expression, Background Emotion × Facial Expression, and Poser Sex × Facial Expression, $BF_{10} = 1.18 \times 10^9$). Comparing the model with the largest BF_{10}, including the three-way Poser Sex × Background Emotion × Facial Expression interaction, to the model with the largest BF_{10}, including the two interactions of Background Emotion × Facial Expression and Poser Sex × Facial Expression, indicated that the model including the two two-way interactions was over 66 times more likely than the model including this three-way interaction, given the observed data.

Discussion

The results of Experiment 1 seem inconsistent with the notion that effects of context scenes on emotion recognition speed are mediated by a match in discrete emotions but rather seem driven by the match between the valence of the scene and the expression. Across the two tasks, the effect of the negative background scenes was uniform, regardless of whether the faces expressed fear or not. However, this conclusion is somewhat tempered by the fact that, unlike the expression of happiness, which was recognized faster than fear or anger on pleasant backgrounds, there was only a trend toward a recognition advantage for the negative expressions on negative backgrounds.

In addition, poser sex affected emotion recognition in that the speed of recognizing happiness was faster than negative expressions for women but not for men's expressions. Finding a larger happy advantage for women compared to men is well established (Lipp et al., 2015). Supplementary Bayesian analyses suggested that an additive model where sex and background emotion independently influenced emotion recognition speed was a better fit for the data than the interactive model (where sex and background emotion interact to influence emotion recognition speed).

Experiment 2

Experiment 1 suggests that background scene valence rather than the emotion match between background scene and facial expression facilitates emotion recognition speed. Poser sex affected emotion recognition speed independently of background scene. Previous studies investigating the influence of social category cues like sex and race on emotional expression have similarly found valence-driven influences in dual-valence tasks (where participants categorize a positive and a negative expression). Interestingly, the more nuanced influences of gender and racial stereotypes, which are not observed in dual-valence tasks, do emerge in single-valence tasks (e.g., where participants categorize two negative expressions; see Bijlstra et al., 2010, 2019). It is possible that both valence match and specific emotion match can lead to context effects, and some of the inconsistencies in the literature have arisen due to valence and specific emotions being confounded in the setup of the experiment.

Experiments 2 and 3 aimed to further investigate specific emotion match effects in single-valence tasks (i.e., where participants encounter only negatively valanced expressions and scenes). If specific emotion match effects are present when valence is held constant, participants should be faster to categorize expressions when face and background are congruent with regards to specific emotion than when face and background are incongruent. Experiments 2 and 3 also aimed to determine whether context effects observed in Experiment 1 generalized beyond the specific face stimuli used in Experiment 1. As such, faces from another database (FACES database; Ebner et al., 2010) were used.

Method

Participants

The approach to power analysis and target sample size reported in Experiment 1 was used for Experiment 2 (target $N = 36$). Due to pandemic conditions, the experiment was conducted online. As online experiments result in higher data loss due to technical problems and reduced participant compliance, 100 positions were available for participation. Rather than terminating the experiment once the quota was hit, students who had already signed up were still given the opportunity to earn course credit. Eighty-three undergraduate students (70 females, 10 males, 3 other, age range: 17–57 years, $M_{age} = 20.07$, $SD_{age} = 5.64$) provided informed consent and provided some response time and/or rating data in exchange for course credit. As expected, there was a higher rate of incomplete data and more errors in the response time task. Twenty-three participants either provided an incomplete data set or had an overall error rate of >20% and were excluded from these analyses. The final sample for the response time task consisted of 60 participants (51 females, 9 males, age range: 17–57, $M_{age} = 20.43$, $SD_{age} = 6.06$). All participants had normal or corrected-to-normal vision. Participants self-reported their ethnicity as Caucasian (43), Asian (8), African (2), Indigenous Australian (1), or other (6). The experimental protocol was approved by the Queensland University of Technology Institutional Ethics Review Board (Human Research Ethics

Application 2022-5019-8378). Experiment 2 was conducted in 2022.

Apparatus, Materials, and Procedure

Participants completed the experiment online on their own desktop computer or laptop. The tasks were controlled using Millisecond's Inquisit 6 Web (Inquisit, 2022). Unlike Javascript and HTML-based online testing platforms that have been associated with more variable and less precise response time measures (e.g., Anwyl-Irvine et al., 2021; Bridges et al., 2020), Inquisit Web 6 installs a small app locally on participants' computers, allowing for precise stimulus presentation timing and response time recording. Pictures of 16 Caucasian posers, eight men and eight women, were drawn from the FACES database (Ebner et al., 2010); poses of fear and disgust were selected for the following models: female: 10, 20, 28, 48, 54, 63, 115, and 152; male: 8, 13, 25, 57, 62, 89, 105, and 153. The images were presented at the center of the screen at 20% of the screen height of the device used by the participant. Four fear-related (1,052, 1,200, 1,300, 1,931) and four disgust-related (7,380, 9,300, 9,570, 9,830) images were selected from the IAPS (Lang et al., 2008) to be employed as emotional background scenes. These images were selected after reviewing Mikels et al.'s (2005) validated ratings of discrete emotional content. Each image was presented simultaneously on the screen four times (in a 2 × 2 grid centered on the middle of the screen) to avoid that the face image obscured the scene content. The grid of contextual images was presented at 60% of the participant's screen size behind the face image. Scenes and face images were presented in color.

Participants completed one categorization task. They were instructed to categorize the facial expression in the center of the screen as fear or disgust by pressing the "S" or "L" keys on their keyboard as quickly and accurately as possible. Response mapping was counterbalanced across participants. Reminders of the response buttons (e.g., "Press S for Disgusted" and "Press L for Fearful") were presented in the bottom left and right corners of the screen for the duration of the task.

The first block of trials consisted of 12 practice trials with response feedback. After completion of the practice trials, participants were presented with two blocks of 128 trials. In total, participants completed 256 trials without response feedback. Faces were divided into two sets (four female and four male faces in each set). Within each block, each of these eight faces that expressed either fear or disgust was combined with each of the eight backgrounds (four fearful, four disgust), resulting in 128 trials. Between experiment blocks, the same backgrounds were used, but the other set of faces was presented matched to backgrounds in the same manner. Block sequence was counterbalanced, and trial sequence was randomized within each block. By the end of the task, each participant was exposed to each face expressing fear and disgust on each background. On each trial, a fixation cross was presented at the center of the screen for 500 ms, followed by the background-face composite image, which remained on the screen until a response was made, or for 3,500 ms. There was a 1,000-ms intertrial interval.

Backgrounds were rated as described in Experiment 1. At this point, demographic information was collected. This information included each participant's age, sex, ethnicity, and the type of device they used to complete the experiment.

Design and Analyses

Data processing proceeded as in Experiment 1. Response times and error rates were submitted to a 2 (Facial Expression: disgust vs. fear) × 2 (Poser Sex: male vs. female) × 2 (Background Emotion: fear vs. disgust) repeated measures ANOVAs. Pairwise comparisons were conducted for all significant interactions. Preliminary analyses revealed no effect of task sequence or participant sex on task performance. As such, these factors are not considered further.

As in Experiment 1, pleasantness ratings for the fear and disgust scenes were averaged within category and subjected to a pairwise t test. Emotion label assignments were added across participants and averaged across images such that a score of 60 would indicate that all participants chose a particular label for a particular set of images (see Supplemental Material for labeling of individual images).

Results

Context Image Ratings and Labeling

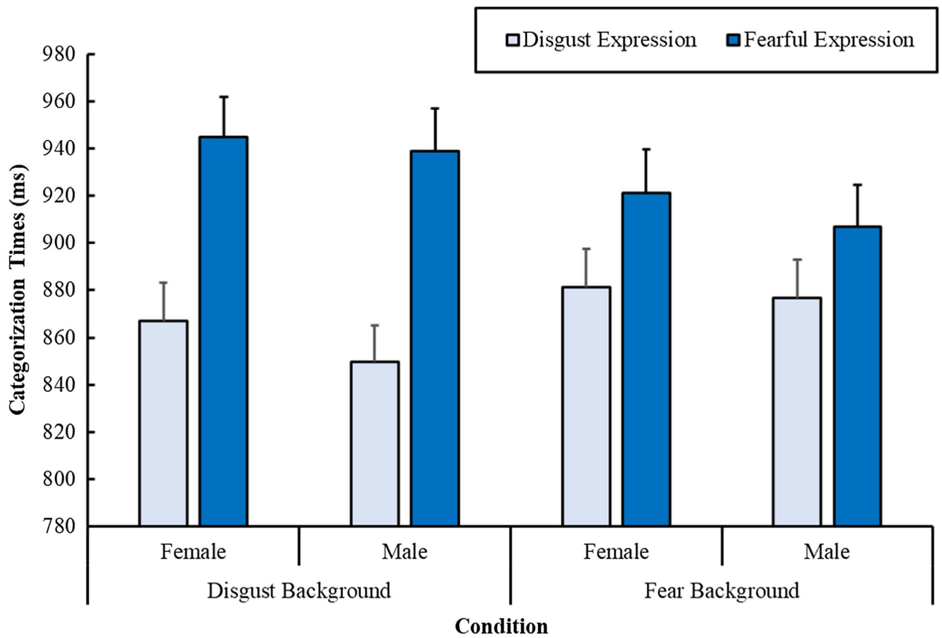
Participants rated the disgust background images ($M = 1.78$, $SD = 0.68$) as less pleasant than the fear backgrounds ($M = 2.87$, $SD = 1.03$), $t(59) = 8.37$, $p < .001$, $d_z = 1.08$. The most common label given to all four fear-eliciting images was "fear" ($M = 45.50$, $SD = 4.45$). The next most common label was surprise ($M = 4.25$, $SD = 2.63$), followed by anger ($M = 3.50$, $SD = 3.87$). The images selected to elicit fear did not elicit sadness, ($M = 0.25$, $SD = 0.50$), contentment, ($M = 2.50$, $SD = 2.38$), happiness, ($M = 1.75$, $SD = 0.96$), or disgust ($M = 2.75$, $SD = 3.59$). Disgust images were most frequently labeled as disgust, with this being the most frequent label given to all four images ($M = 46.00$, $SD = 13.24$). One of the disgust images (IAPS 9,470) was also frequently interpreted as sad (22 participants vs. 29 participants who selected disgust). On average, sad was the next most common emotion label ($M = 6.25$, $SD = 10.59$), followed by anger ($M = 3.25$, $SD = 3.95$). Images were not interpreted as contentment ($M = 0.75$, $SD = 0.96$), happiness ($M = 0.25$, $SD = 0.50$), fear ($M = 2.75$, $SD = 0.50$), or surprise ($M = 0.75$, $SD = 0.96$). These results are consistent with the data provided by Mikels et al. (2005).

Emotion Recognition Speed

As shown in Figure 2, consistent with the specific emotion match account, background emotion affected the recognition speed of fear and disgust expressions, such that participants were faster to recognize both fearful and disgust expressions when the background was emotionally congruent than when the background was incongruent. Overall, participants were also faster to recognize disgust than fear expressions, $F(1, 59) = 34.51$, $p < .001$, $\eta_p^2 = .37$. Consistent with a context effect, there was a significant interaction between facial expression and background emotion, $F(1, 59) = 11.74$, $p = .001$, $\eta_p^2 = .17$. Comparisons conducted to follow up on this significant interaction indicated that participants were faster to categorize disgust faces in both disgust, $F(1, 59) = 38.34$, $p < .001$, $\eta_p^2 = .39$, and fearful backgrounds, $F(1, 59) = 10.04$, $p = .002$, $\eta_p^2 = .15$, but the effect was significantly larger in disgust backgrounds. No other effects were significant, $F < 1.87$, $p > .177$, $\eta_p^2 < .03$.

Figure 2

Recognition Times for Disgust and Fear Expressions on Women's and Men's Faces as a Function of Background Scene Emotion and in Experiment 2 (Error Bars Represent Standard Errors of the Mean)



Note. See the online article for the color version of this figure.

Supplementary Bayesian Analyses

Consistent with the frequentist approach, the model with the largest BF was the model including the main effects of facial expression and background and the Facial Expression × Background Emotion interaction, $BF_{10} = 7.84 \times 10^{11}$. The model with the largest BF including Poser Sex × Facial Expression and Background Emotion × Facial Expression interactions, $BF_{10} = 3.06 \times 10^{10}$, was 29.82 times more likely than the model, including the three-way Poser Sex × Facial Expression × Background Emotion interaction given the observed data.

Discussion

As predicted, consistent with the specific emotion match explanation, a background context effect was observed in the context of a single-valence task. Participants were faster to categorize expressions on emotionally congruent backgrounds than incongruent backgrounds. These results suggest that specific emotion congruence effects can be observed in contexts where broad valence is not task-relevant. Unlike Experiment 1, however, Experiment 2 found no influence of face sex on emotion categorization. Although fear expressions have been found to be more stereotypically associated with women than men (Plant et al., 2000), there was no influence of poser sex on emotion categorization speeds in this experiment. This may suggest that when only fear and disgust expressions are encountered in a task, no strong automatic stereotypic association between men or women and fear or disgust was elicited. Another possibility is that the inclusion of negative background scenes reduced the influence of gender stereotypes on emotion recognition previously observed in single-valence tasks (Bijlstra et al., 2010). The failure to find a poser sex effect on emotion categorization is not due to the relatively lower percentage of male participants in the sample of Experiment 2. The poser sex effects found in Experiment 1 (and Experiment 3—see the following section) are still evident if data from female participants only were analyzed.

Experiment 3

The aim of Experiment 3 was to replicate and extend the specific emotion-match context effect observed in Experiment 2 beyond the specific backgrounds and emotions used in Experiment 2. Selecting a different combination of emotions also provides the opportunity to further explore the interaction between sex and emotion in a task where background contexts are also varied to confirm the finding of Experiment 1 that Poser Sex × Facial Expression effects are independent of background effects. As such, in Experiment 3, participants categorized fearful and sad expressions placed on backgrounds eliciting fear or sadness. Sadness is an ideal negative emotion to investigate further, as it is stereotypically linked to women (Plant et al., 2000). Sadness has also been used successfully in past research investigating the influence of gender stereotypes on emotion recognition (Bijlstra et al., 2010).

Method

Participants

Sample size was determined as for Experiment 2; however, as this experiment was conducted toward the end of the semester and we did not want to prevent participants from obtaining their course credit, we did not close recruitment once the target sample size was reached. A total of 114 participants gave informed consent and provided some response time and ratings data for the current experiment (83 female, 27 male, 2 nonbinary, 2 chose not to disclose, age range: 16–52 years, $M_{\text{age}} = 21.65$, $SD_{\text{age}} = 6.39$). The majority of the sample identified as Caucasian (80), followed by Asian (16), African (4), Indigenous Australian (4), or other (7). Three participants chose not to disclose their ethnicity. After removing participants who provided incomplete data sets or who had error rates >20%, 97 participants' data were included in the final analyses (age range: 16–52, $M_{\text{age}} = 22.02$, $SD_{\text{age}} = 6.77$). There were 72 women, 23 men, one nonbinary person, and one person who did not disclose their sex. The majority of participants identified as Caucasian (67), followed by Asian (13), African (4), and Indigenous Australian (4). Seven participants nominated another ethnicity, and two did not disclose their ethnicity. The procedure for Experiment 3 was approved by the Queensland University of Technology Institutional Ethics Review Board (HREA 2022-5019-8378). Experiment 3 was conducted in 2022.

Apparatus, Materials, and Procedure

Experiment 3 proceeded as did Experiment 2 with the following differences. The same posers were presented; however, disgust expressions were replaced with each poser's sad expression. The four disgust scenes were replaced with four scenes eliciting sadness (2,053, 3,230, 3,350, 9,000), and four new fear scenes were selected (1,110, 1,113, 1,301, 1,930) from the IAPS (Lang et al., 2008) to be employed as emotional background scenes. These images were selected after reviewing Mikels et al.'s (2005) validated ratings of discrete emotional content.

Design and Analyses

Data processing proceeded as in Experiment 2. Response times and error rates were submitted to a 2 (Facial Expression: sad vs. fear) × 2 (Poser Sex: male vs. female) × 2 (Background Emotion: fear vs. sadness) repeated measures ANOVAs. Pairwise comparisons were conducted for all significant interactions. As preliminary analyses revealed no effect of task sequence or participant sex on task performance, these factors are not considered further.

As in Experiment 1, pleasantness ratings for the fearful and sad scenes were averaged within category and subjected to a pairwise *t* test. Emotion label assignments were added across participants and averaged across images such that a score of 97 would indicate that all participants chose a particular label for a particular set of images (see Supplemental Material for labeling of individual images).

Results

Context Image Ratings and Labeling

Participants rated the sad background images ($M = 2.14$, $SD = 0.72$) as less pleasant than the fear backgrounds ($M = 2.72$, $SD = 0.86$), $t(96) = 7.83$, $p < .001$, $d_z = 0.80$. The most common label given to all four fear-related images was "fear" ($M = 67.00$, $SD = 12.49$). The next most common emotion ascribed was surprise ($M = 15.25$, $SD = 3.69$), followed by contentment ($M = 7.00$, $SD = 6.93$). Images were not recognized as conveying sadness ($M = 1.00$, $SD = 1.15$), happiness ($M = 0.75$, $SD = 1.5$), anger ($M = 3.75$, $SD = 4.99$), or disgust ($M = 2.25$, $SD = 1.71$). All four sad images were most frequently labeled as sad ($M = 75.25$, $SD = 19.29$). One of the sad images (9,000—cemetery) was also frequently interpreted as fear (23 participants) or contentment (21 participants) compared to 47 who selected sad. So, on average, fear was the next most common emotion ascribed ($M = 9.50$, $SD = 9.14$), followed by contentment ($M = 6.25$, $SD = 9.81$). Images were not recognized as eliciting happiness ($M = 1.75$, $SD = 1.50$), surprise ($M = 1.00$, $SD = 0.82$), anger ($M = 0.50$, $SD = 0.58$), or disgust ($M = 2.50$, $SD = 2.08$). These results are consistent with the data provided by Mikels et al. (2005).

Emotion Recognition Speed

As shown in Figure 3, consistent with the specific emotion match account, background emotion affected the speed of recognizing fearful and sad expressions, such that participants were faster to recognize both fearful and sad expressions when the background was emotionally congruent than when the background was incongruent. Additionally, the sex of the face moderated the speed of recognizing fearful versus sad expressions.

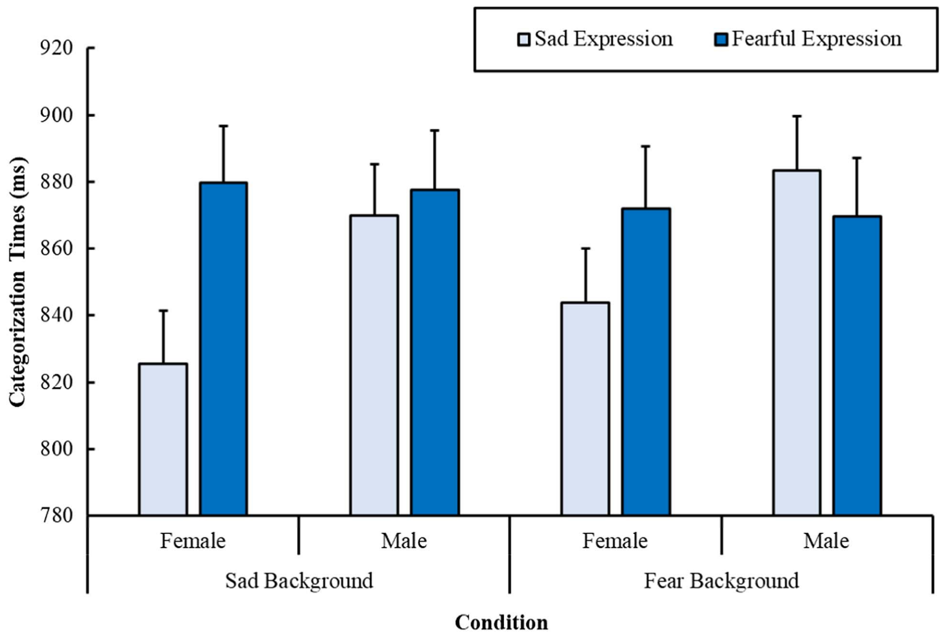
Looking at the results of the repeated measures ANOVA, there was a main effect of facial expression such that participants were faster to recognize sad than fearful expressions, $F(1, 96) = 5.98$, $p = .016$, $\eta_p^2 = .06$. A main effect of poser sex also emerged, such that expressions were recognized faster on female than on male faces, $F(1, 96) = 16.29$, $p < .001$, $\eta_p^2 = .15$. These effects were moderated by a significant Poser Sex × Facial Expression interaction, $F(1, 96) = 15.76$, $p < .001$, $\eta_p^2 = .14$, which emerged as participants were significantly faster to categorize sad compared to fearful expressions on female faces, $F(1, 96) = 20.33$, $p < .001$, $\eta_p^2 = .18$, but not on male faces, $F(1, 96) = 0.10$, $p = .759$, $\eta_p^2 < .01$. Consistent with the emotion match account, there was a significant Background Emotion × Emotional Expression interaction, $F(1, 96) = 11.22$, $p = .001$, $\eta_p^2 = .10$. Follow-up comparisons indicated that participants were faster to categorize sadness than fear on sadness eliciting backgrounds, $F(1, 96) = 11.26$, $p = .001$, $\eta_p^2 = .11$, but not on fear eliciting backgrounds, $F(1, 96) = 0.83$, $p = .364$, $\eta_p^2 = .01$. No other effects were significant, $F < 0.80$, $p > .372$, $\eta_p^2 < .01$.

Supplementary Bayesian Analyses

The model with the largest BF was the one including the main effects of facial expression and poser sex and the Facial Expression × Poser Sex interaction, $BF_{10} = 2.16 \times 10^6$. The next strongest model included the main effects of background emotion, poser sex, and facial expression, as well as Background Emotion × Facial Expression and Poser Sex × Facial Expression interactions, $BF_{10} = 334263.79$. Comparing this additive model to the model including the three-way Poser Sex × Background Emotion × Facial Expression interaction ($BF_{10} = 5546.18$), it was indicated that the model including the two-way Poser Sex × Facial Expression and Background Emotion × Facial Expression interactions was 60.26 times more likely than the model including the three-way interaction given the observed data.

Figure 3

Recognition Times for Fearful and Sad Expressions on Women's and Men's Faces as a Function of Background Scene Emotion and in Experiment 3 (Error Bars Represent Standard Errors of the Mean)



Note. See the online article for the color version of this figure.

Discussion

As predicted, consistent with Experiment 2, a specific emotion match effect was observed in Experiment 3. When categorizing fearful and sad expressions placed on fear and sadness-related backgrounds, participants were faster to recognize sadness when face and background were congruent than when they were incongruent. Unlike in Experiment 2, but consistent with Experiment 1, poser sex also independently moderated the speed of emotion recognition. Participants were significantly faster to categorize sad than fearful expressions on female faces, but not on male faces.

General Discussion

The aims of the present study were twofold. First, this study aimed to investigate whether the broad valence or specific emotion elicited by background scenes drove the influence of contextual scenes on emotion recognition speed. It was predicted that if facilitation of emotion recognition speed is limited to discrete emotion matches, then faster recognition should be evident only for expressions that specifically match the emotion elicited by the background. If it is scene valence match that facilitates emotion recognition, then the same background context effects should be evident for both fear and anger. The results of Experiment 1 demonstrate that broad valence match, and not specific emotion match, drives context effects in tasks where broad positive–negative valence is task-relevant (i.e., in a dual-valence task). Experiments 2 and 3 demonstrate that specific emotion match does drive context effects where broad valence is not relevant to the participant's task (i.e., in a single-valence task).

The second aim of this study was to clarify how background scene contexts and social information on the face (i.e., poser sex) may combine to influence emotion perception. Although many past studies have included female and male posers, none had included this variable in analysis. In Experiments 1 and 3, emotion recognition speed was affected by poser sex. In Experiment 1, happiness was detected faster than anger or fear when expressed by women but not when expressed by men. In Experiment 3, where only negative expressions were presented, participants were faster to categorize sadness than fear on female faces, but not male faces. In both cases, there was no higher order interaction of poser sex, background context, and emotional expression, suggesting that poser sex and background context had independent influences on emotion recognition speed. Supplementary Bayesian analyses supported this interpretation.

Considering the findings within the broader literature, the finding that broad valence, but not specific emotion match, influenced the speed of emotion recognition where broad valence is task-relevant is consistent with a number of previous studies investigating context effects with both positive and negative expressions (e.g., Hietanen & Astikainen, 2013; Zhang et al., 2013). To our knowledge, Experiments 2 and 3 were the first to investigate context effects using a single-valence task design. Consistent with the broader literature investigating the influence of the stereotypic associations between face sex and ethnicity on emotion perception (Bijlstra et al., 2010), specific emotion effects are not observed in dual-valence tasks but emerge in a single-valence task once the evaluative dimension is no longer made salient.

Implications for Theory

These results suggest a possible hierarchy of influences of both social and scene-based contextual factors. Findings suggest that broad positive–negative evaluations of scenes and social cues drive the influence of contextual information on emotion recognition speed over and above specific emotion or stereotype congruence in situations where both positive and negative expressions are encountered. The influence of congruence between specific negative emotions arises once the valence is no longer task-relevant. These findings are consistent with the idea that both contextual scenes and social stimuli like faces are processed first at the global level (based on broad valence) before being processed at the specific level (i.e., distinguishing between specific negative emotions). The idea that emotional reactions progress from global (broad evaluation) to specific has been promoted by Ruys and Stapel (e.g., see Ruys & Stapel, 2008, though empirical evidence was subsequently retracted due to fraudulent data; Ruys & Stapel, 2012). The idea that evaluation occurs early in the processing of and response to emotional stimuli prior to differentiation of specific emotions is also consistent with certain appraisal theories like the multicomponent process model (Scherer & Moors, 2019). It is also consistent with the idea that basic affective responding is a core process preceding the construction of the meaning of specific emotions within constructionist theories of emotion (e.g., Barrett, 2006). The idea of primacy of valence before specific emotion when it comes to the influence of contextual information on emotion recognition speed is borne out in these data.

The finding that poser sex independently influenced emotion perception along with background context also replicates and extends findings from the previous literature. Only a few studies have focused on investigating the concurrent influence of multiple sources of contextual information on emotion perception, mainly focusing on the influence of multiple social cues in the face (e.g., race and sex, age and sex; Craig & Lipp, 2018; Lindeberg et al., 2019; Smith et al., 2017; Tipples, 2023). Supplementary Bayesian analyses conducted for the current research suggest that background scenes and poser sex have an additive influence, both independently influencing the speed of recognizing emotional expressions. This is consistent with past studies investigating how multiple concurrently present social cues (i.e., race and sex, or age and sex) influence emotion recognition speed and accuracy (Craig & Lipp, 2018).

Limitations

The findings of the present study must be considered within the context of certain limitations. Background images were selected with the goal of eliciting specific emotions, though there was some variability in the interpretation of emotional content from the background. Overall, the majority of participants interpreted the emotional content of the image in line with the intended emotion category, but there were some participants who interpreted the image as representative of a different specific emotion or sometimes even a different valence altogether. Although this variability is unlikely to pose an alternative explanation for the results, where participants construed the background as having a different meaning than intended, this may have added noise to the influence of contexts on face recognition speed.

While Experiment 1 was conducted in the laboratory, Experiments 2 and 3 were conducted online. Conducting experiments online comes with increased ease in accessing participants, but also an increase in participants who provide incomplete data sets. This can be observed when comparing the number of excluded participants in Experiment 1 versus Experiments 2 and 3. Conducting experiments online also comes with less control over the testing environment. Participants complete the experiment in an environment uncontrolled by the experimenter, and screen size and resolution will vary. Although only correct response times are included in analyses and participants with very high error rates suggesting inattention to the task were excluded from analysis, a lack of experimental control may mean participants are distracted and less attentive. Variance in viewing conditions (lighting, screen size, resolution, etc.) may also introduce greater statistical noise to response time analyses. On the other hand, the ease of recruiting larger sample sizes can offset some of the downsides of online data collection. Comparing Experiment 1 to Experiments 2 and 3 suggests comparable variance in response times in laboratory and online environments.

Constraints on Generality

The current research was conducted using samples recruited from undergraduate courses at two large Australian universities. As such, the samples comprise a majority of women who declared their ethnicity as Caucasian. This can impose constraints on the generality of our findings as the samples are on average younger and better educated than the general population and overrepresent Caucasian women. Gender imbalance may impose a limitation, although our past work has not revealed evidence for effects of participant gender when assessed in preliminary analyses, and future research should extend the work to more balanced samples with a larger range of ages and ethnicities. Having a better representation of male participants would also permit the explicit assessment of participant sex effects, which is not possible given the small number of men included.

Conclusion

In summary, the current research advances our understanding of the manner in which facial expressions are recognized in a number of ways. First, it suggests that the effect of emotional scenes on the speed of recognizing facial expressions of emotion can be mediated by a match in stimulus valence and that a match of the discrete emotion related to the scene and displayed by the face does not seem to provide for an additional recognition advantage in tasks where positive and negative expressions are encountered. This is consistent with Hugenberg's (2005) evaluative congruence account and complements the research on the effects of facial social category cues on emotion recognition. Specific emotion match effects (as well as the influence of sex stereotypes) can be observed in tasks where only negative expressions are categorized. Finally, the current research provided additional evidence that multiple concurrently present sources of information do affect emotion recognition simultaneously and, at least in the case of emotional scenes and facial sex cues, independently.

References

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*, *53*(4), 1407–1425. https://doi.org/10.3758/s13428-020-01501-5

Barrett, L. F. (2006). Are emotions natural kinds? *Perspectives on Psychological Science*, *1*(1), 28–58. https://doi.org/10.1111/j.1745-6916.2006.00003.x

Beaupré, M. G., & Hess, U. (2005). Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology*, *36*(3), 355–370. https://doi.org/10.1177/0022022104273656

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, *92*(2), 179–190. https://doi.org/10.1037/0022-3514.92.2.179

Bijlstra, G., Holland, R. W., Dotsch, R., & Wigboldus, D. H. J. (2019). Stereotypes and prejudice affect the recognition of emotional body postures. *Emotion*, *19*(2), 189–199. https://doi.org/10.1037/emo0000438

Bijlstra, G., Holland, R. W., & Wigboldus, D. H. J. (2010). The social face of emotion recognition: Evaluations versus stereotypes. *Journal of Experimental Social Psychology*, *46*(4), 657–663. https://doi.org/10.1016/j.jesp.2010.03.006

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, Article e9414. https://doi.org/10.7717/peerj.9414

Bryce, L., Mika, G., Craig, B., Hess, U., & Lipp, O. V. (2023). *Emotional scenes as context in emotional expression recognition*. https://osf.io/4h6v7/

Craig, B. M., Koch, S., & Lipp, O. V. (2017). The influence of social category cues on the happy categorisation advantage depends on expression valence. *Cognition and Emotion*, *31*(7), 1493–1501. https://doi.org/10.1080/02699931.2016.1215293

Craig, B. M., & Lee, A. J. (2020). Stereotypes and structure in the Interaction between facial emotional expression and sex characteristics. *Adaptive Human Behavior and Physiology*, *6*(2), 212–235. https://doi.org/10.1007/s40750-020-00141-5

Craig, B. M., & Lipp, O. V. (2017). The influence of facial sex cues on emotional expression categorization is not fixed. *Emotion*, *17*(1), 28–39. https://doi.org/10.1037/emo0000208

Craig, B. M., & Lipp, O. V. (2018). The influence of multiple social categories on emotion perception. *Journal of Experimental Social Psychology*, *75*, 27–35. https://doi.org/10.1016/j.jesp.2017.11.002

Craig, B. M., Zhang, J., & Lipp, O. V. (2017). Facial race and sex cues have a comparable influence on emotion recognition in Chinese and Australian participants. *Attention, Perception, & Psychophysics*, *79*(7), 2212–2223. https://doi.org/10.3758/s13414-017-1364-z

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*(3), 289–311. https://doi.org/10.1080/026999300378824

Diéguez-Risco, T., Aguado, L., Albert, J., & Hinojosa, J. A. (2013). Faces in context: Modulation of expression processing by situational information. *Social Neuroscience*, *8*(6), 601–620. https://doi.org/10.1080/17470919.2013.834842

Eagly, A. H., Mladinic, A., & Otto, S. (1991). Are women evaluated more favorably than men? An analysis of attitudes, beliefs, and emotions. *Psychology of Women Quarterly*, *15*(2), 203–216. https://doi.org/10.1111/j.1471-6402.1991.tb00792.x

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, *42*(1), 351–362. https://doi.org/10.3758/BRM.42.1.351

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Hess, U., Adams, R. B., Jr., & Kleck, R. E. (2009). The face is not an empty canvas: How facial expressions interact with facial appearance. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1535), 3497–3504. https://doi.org/10.1098/rstb.2009.0165

Hess, U., & Hareli, S. (2016). The impact of context on the perception of emotions. In C. Abell & J. Smith (Eds.), *The expression of emotion: Philosophical, psychological, and legal perspectives* (pp. 199–218). Cambridge University Press. https://doi.org/10.1017/CBO9781316275672.010

Hietanen, J. K., & Astikainen, P. (2013). N170 response to facial expressions is modulated by the affective congruency between the emotional expression and preceding affective picture. *Biological Psychology*, *92*(2), 114–124. https://doi.org/10.1016/j.biopsycho.2012.10.005

Hugenberg, K. (2005). Social categorization and the perception of facial affect: Target race moderates the response latency advantage for happy faces. *Emotion*, *5*(3), 267–276. https://doi.org/10.1037/1528-3542.5.3.267

Hugenberg, K., & Sczesny, S. (2006). On wonderful women and seeing smiles: Social categorization moderates the happy face response latency advantage. *Social Cognition*, *24*(5), 516–539. https://doi.org/10.1521/soco.2006.24.5.516

Inquisit. (2015). (Version 4.1) [Computer software]. https://www.millisecond.com

Inquisit. (2022). (Version 6.61) [Computer software]. https://www.millisecond.com

JASP Team. (2024). *JASP* (Version 0.18.3) [Computer software]. https://jasp-stats.org/download/

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kastendieck, T., Mauersberger, H., Blaison, C., Ghalib, J., & Hess, U. (2021). Laughing at funerals and frowning at weddings: Top-down influences of context-driven social judgments on emotional mimicry. *Acta Psychologica*, *212*, Article 103195. https://doi.org/10.1016/j.actpsy.2020.103195

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual* [Technical Report A-8]. University of Florida.

Leppänen, J. M., & Hietanen, J. K. (2003). Affect and face perception: Odors modulate the recognition advantage of happy faces. *Emotion*, *3*(4), 315–326. https://doi.org/10.1037/1528-3542.3.4.315

Lindeberg, S., Craig, B. M., & Lipp, O. V. (2019). 2:0 for the good guys: Character information influences emotion perception. *Emotion*, *19*(8), 1495–1499. https://doi.org/10.1037/emo0000530

Lipp, O. V., Craig, B. M., & Dat, M. C. (2015). A happy face advantage with male Caucasian faces: It depends on the company you keep. *Social Psychological and Personality Science*, *6*(1), 109–115. https://doi.org/10.1177/1948550614546047

Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(45), 16518–16523. https://doi.org/10.1073/pnas.0507650102

Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., & Reuter-Lorenz, P. A. (2005). Emotional category data on images from the International Affective Picture System. *Behavior Research Methods*, *37*(4), 626–630. https://doi.org/10.3758/BF03192732

Milanak, M. E., & Berenbaum, H. (2014). The effects of context on facial affect recognition. *Motivation and Emotion*, *38*(4), 560–568. https://doi.org/10.1007/s11031-014-9401-x

Morey, R. D., Romeijn, J.-W., & Rouder, J. N. (2016). The philosophy of Bayes factors and the quantification of statistical evidence. *Journal of*

Mathematical Psychology, *72*, 6–18. https://doi.org/10.1016/j.jmp.2015.11.001

Niedenthal, P. M., & Brauer, M. (2012). Social functionality of human emotion. *Annual Review of Psychology*, *63*(1), 259–285. https://doi.org/10.1146/annurev.psych.121208.131605

Plant, E. A., Hyde, J. S., Keltner, D., & Devine, P. G. (2000). The gender stereotyping of emotions. *Psychology of Women Quarterly*, *24*(1), 81–92. https://doi.org/10.1111/j.1471-6402.2000.tb01024.x

Righart, R., & de Gelder, B. (2008a). Rapid influence of emotional scenes on encoding of facial expressions: An ERP study. *Social Cognitive and Affective Neuroscience*, *3*(3), 270–278. https://doi.org/10.1093/scan/nsn021

Righart, R., & de Gelder, B. (2008b). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective & Behavioral Neuroscience*, *8*(3), 264–272. https://doi.org/10.3758/CABN.8.3.264

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Ruys, K. I., & Stapel, D. A. (2008). How to heat up from the cold: Examining the preconditions for (unconscious) mood effects. *Journal of Personality and Social Psychology*, *94*(5), 777–791. https://doi.org/10.1037/0022-3514.94.5.777

Ruys, K. I., & Stapel, D. A. (2012). “How to heat up from the cold: Examining the preconditions for (unconscious) mood effects”: Retraction of Ruys and Stapel (2008). *Journal of Personality and Social Psychology*, *103*(3), Article 451. https://doi.org/10.1037/a0029740

Scherer, K. R., & Moors, A. (2019). The emotion process: Event appraisal and component differentiation. *Annual Review of Psychology*, *70*(1), 719–745. https://doi.org/10.1146/annurev-psych-122216-011854

Smith, J. S., LaFrance, M., & Dovidio, J. F. (2017). Categorising intersectional targets: An “either/and” approach to race- and gender-emotion congruity. *Cognition and Emotion*, *31*(1), 83–97. https://doi.org/10.1080/02699931.2015.1081875

Tipples, J. (2023). When men are wonderful: A larger happy face facilitation effect for male (vs. female) faces for male participants. *Emotion*, *23*(7), 2080–2093. https://doi.org/10.1037/emo0001221

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Wieser, M. J., & Brosch, T. (2012). Faces in context: A review and systematization of contextual influences on affective face processing. *Frontiers in Psychology*, *3*, Article 471. https://doi.org/10.3389/fpsyg.2012.00471

Xu, Q., Yang, Y., Tan, Q., & Zhang, L. (2017). Facial expressions in context: Electrophysiological correlates of the emotional congruency of facial expressions and background scenes. *Frontiers in Psychology*, *8*, Article 2175. https://doi.org/10.3389/fpsyg.2017.02175

Xu, Q., Yang, Y., Zhang, E., Qiao, F., Lin, W., & Liang, N. (2015). Emotional conflict in facial expression processing during scene viewing: An ERP study. *Brain Research*, *1608*, 138–146. https://doi.org/10.1016/j.brainres.2015.02.047

Zhang, X., Liu, Z., & Xuan, Y. (2013). *The role of emotional scenes in facial expressional recognition: Priming versus contextual effect* [Conference session]. Proceedings of Third International Conference on Information Science and Technology, Yangzhou, Jiangsu, China. https://doi.org/10.1109/ICIST.2013.6747489

Received March 4, 2021
Revision received October 30, 2023
Accepted February 2, 2024 ■